

## 明 細 書

## マルチチャネル情報処理装置

5

## (技術分野)

本発明は、音声データを伴う複数の動画像情報を同時に表示装置上に表示可能なマルチチャネル情報処理に関し、特に、表示装置上に表示された複数の動画像の中から特定の動画像の内容を容易に把握することができるマルチチャネル情報

10 処理装置、マルチチャネル情報処理方法、マルチチャネル情報処理方法のプログラムおよびこのプログラムを記憶するコンピュータ読み取り可能な記憶媒体に関する。

## (背景技術)

15 近年、ADSL (Asymmetric Digital Subscriber Line) やCATVモデムなどのインターネットに高速にアクセスする手段が普及しており、これに伴ってTV番組などの動画像情報を家庭からインターネットを通じてダウンロードする機会が増えている。また、デジタルビデオカメラとパーソナルコンピュータ (以下、PCと称す) との間のインターフェイスが普及しており、ユーザが自分で撮

20 影した動画像情報をPCに取り込んで処理することも盛んに行われている。

また、PCの性能が向上していることやハードディスクの容量が拡大していることから、大量の動画像情報をハードディスクに格納して再生表示することも容易になってきており、複数の動画像情報を表示装置上に同時に表示できる環境にも容易に対応可能となっている。このような複数の動画像情報の表示には、CR

25 Tや液晶ディスプレイなどの表示装置上に複数のウィンドウを任意の位置に配置し、これらの動画像情報を同時に表示させることが可能である。

このような複数の動画像を表示する場合に、これら動画像に付随している音声データは、全ての音声データを出力しないか、あるいは全ての音声データを同一の音量で合成して出力することが行われている。音声データを出力しない場合に

は、各動画像情報の内容を把握するために音声データを利用することができず、表示装置上に表示された画像のみでその内容を判断する必要がある。また、全ての音声データを同一の音量で合成して出力する場合には、それぞれの音声データが干渉しあって音声を認識することが困難であり、またどの音声データがどの動画像情報に対応しているかを認識することも困難であるという問題を包含している。

本発明は、複数の動画像情報を表示装置上に表示する際に、特定の動画像情報に対応する音声データを認識することを容易にし、表示されている動画像情報の内容を容易に把握することを可能にすることを目的とする。

10

#### (発明の開示)

本発明の請求項 1 に係るマルチチャネル情報処理装置は、複数の動画像情報を同時に表示装置に表示するマルチチャネル情報処理装置であって、複数の動画像情報を取得し、複数の動画像情報の表示装置上での表示位置に関する動画像位置情報を決定して複数の動画像情報を前記動画像位置情報に基づいて出力する動画像情報制御手段と、入力デバイスを介して入力されるカーソル指示情報に基づいてカーソル位置情報を算出しカーソル位置情報に基づいてカーソル画像情報を生成してこれを出力するカーソル位置制御手段と、動画像表示制御手段が出力する複数の動画像情報とカーソル位置制御手段が出力するカーソル画像情報とを合成して表示装置上に表示する表示画像生成手段と、複数の動画像情報に対応する動画像位置情報とカーソル位置制御手段で算出されたカーソル位置情報とに基づいて、各動画像情報の表示位置とカーソルの表示位置との距離を算出し距離情報を生成する距離情報生成手段と、距離情報生成手段で生成された距離情報に基づいて複数の動画像情報に対応する音声データの音量を決定して音声出力装置に出力する音声出力制御手段とを備える。

25

本発明の請求項 2 に係るマルチチャネル情報処理装置は、請求項 1 に記載のマルチチャネル情報処理装置であって、音声出力制御手段が、距離情報生成手段で生成された距離情報のうち距離の値が最も小さい動画像情報に対応する音声デ

タの音量を予め定められた音量 $V_a$ に設定し、他の動画像情報に対応する音声データの音量を音量 $V_a$ よりも小さな予め定められた音量 $V_b$ に設定することを特徴としている。

- 5      本発明の請求項 3 に係るマルチチャネル情報処理装置は、請求項 1 に記載のマルチチャネル情報処理装置であって、音声出力制御手段が、複数の動画像情報に対応する音声データの音量を、距離情報生成手段で生成された距離情報中の距離の値に反比例するように設定することを特徴としている。
- 10      本発明の請求項 4 に係るマルチチャネル情報処理装置は、請求項 1 に記載のマルチチャネル情報処理装置であって、距離情報生成手段で生成された距離情報は、カーソルの表示位置から見た各動画像情報の表示位置の方向に関する方向情報を含み、音声出力制御手段は音声出力装置が形成する音像空間に複数の動画像情報に対応する音声データが定位するように距離情報に基づいて音声出力装置に出力
- 15      を行うことを特徴としている。

- 本発明の請求項 5 に係るマルチチャネル情報処理装置は、請求項 1 に記載のマルチチャネル情報処理装置であって、複数の動画像情報に対応する音声データに含まれる言葉を認識する音声データ認識手段と、音声データ認識手段で認識した
- 20      言葉を文字情報に変換して表示装置に表示する文字情報表示手段とをさらに備えている。

- 本発明の請求項 6 に係るマルチチャネル情報処理装置は、請求項 5 に記載のマルチチャネル情報処理装置であって、インターネット接続手段と、音声データ認識
- 25      手段で認識した言葉をキーワードとしてインターネット上に存在する関連ホームページを検索するホームページ検索手段と、ホームページ検索手段で検索されたホームページを前記表示装置上に表示するホームページ表示手段とを備えている。

本発明の請求項 7 に係るマルチチャネル情報処理装置は、請求項 1 に記載のマルチチャネル情報処理装置であって、表示装置に表示される複数の動画像情報から所定のアルゴリズムに基づいて特定の動画像情報を選択する動画像選択手段をさらに備え、音声出力制御手段は動画像選択手段が選択した動画像情報に対応する音声データを音声出力装置に出力することを特徴としている。

本発明の請求項 8 に係るマルチチャネル情報処理装置は、請求項 7 に記載のマルチチャネル情報処理装置であって、動画像選択手段は、所定時間毎に選択する動画像情報を切り換えることを特徴としている。

10

本発明の請求項 9 に係るマルチチャネル情報処理方法は、複数の動画像情報を同時に表示装置に表示する際のマルチチャネル情報処理方法であって、表示装置上に表示を行う動画像情報表示装置上での表示位置を決定する段階と、決定した表示位置に基づいて複数の動画像情報を出力する段階と、入力デバイスから入力されるカーソル指示情報を受け付ける段階と、カーソル指示情報に基づいてカーソルを表示するカーソル位置情報を算出する段階と、カーソル位置情報に基づいてカーソル画像情報を生成する段階と、複数の動画像情報とカーソル画像情報とを合成して表示画像を生成しこれを表示装置上に出力する段階と、複数の動画像情報の表示位置とカーソル位置情報との間の距離を算出し距離情報を生成する段階と、距離情報に基づいて複数の動画像情報に対応する音声データの音量を決定し音声出力装置に出力する段階とを含んでいる。

20

本発明の請求項 10 に係るマルチチャネル情報処理方法は、請求項 9 に記載のマルチチャネル情報処理方法であって、複数の動画像情報について生成された距離情報のうち距離の値が最も小さい動画像情報に対応する音声データの音量を予め定められた音量  $V_a$  に設定し、他の動画像情報に対応する音声データの音量を音量  $V_a$  よりも小さな予め定められた音量  $V_b$  に設定することを特徴としている。

25

本発明の請求項 1 1 に係るマルチチャネル情報処理方法は、請求項 9 に記載のマルチチャネル情報処理方法であって、複数の動画像情報に対応する音声データの音量を、各距離情報中の距離の値に反比例するように設定することを特徴としている。

5

本発明の請求項 1 2 に係るマルチチャネル情報処理方法は、請求項 9 に記載のマルチチャネル情報処理方法であって、カーソルの表示位置から見た各動画像情報の表示位置の方向に関する方向情報を生成する段階と、複数の動画像情報に対応する音声データが、音声出力装置の音像空間における距離情報および方向情報  
10 に応じた音像位置に定位するように、音声出力装置に出力を行う段階とをさらに備えている。

本発明の請求項 1 3 に係るマルチチャネル情報処理方法は、請求項 9 に記載のマルチチャネル情報処理方法であって、複数の動画像情報に対応する音声データ  
15 に含まれる言葉を音声認識する段階と、音声認識した言葉を文字情報に変換して表示装置に表示する段階とをさらに備えている。

本発明の請求項 1 4 に係るマルチチャネル情報処理方法は、請求項 1 3 に記載のマルチチャネル情報処理方法であって、インターネットに接続する段階と、音声認識した言葉をキーワードとしてインターネット上に存在する関連ホームページを検索する段階と、検索された関連ホームページを前記表示装置上に表示する  
20 段階とを備えている。

本発明の請求項 1 5 に係るマルチチャネル情報処理方法は、請求項 9 に記載の  
25 マルチチャネル情報処理方法であって、表示装置に表示される複数の動画像情報から所定のアルゴリズムに基づいて特定の動画像情報を選択する段階と、選択した動画像情報に対応する音声データを音声出力装置に出力する段階とをさらに備えている。

本発明の請求項 16 に係るプログラムは、複数の動画像情報を同時に表示装置に表示する際のマルチチャネル情報処理方法のプログラムであって、複数の動画像情報の表示装置上での表示位置を決定する段階と、決定した表示位置に基づいて複数の動画像情報を出力する段階と、入力デバイスから入力されるカーソル指示情報を受け付ける段階と、カーソル指示情報に基づいてカーソルを表示するカーソル位置情報を算出する段階と、カーソル位置情報に基づいてカーソル画像情報を生成する段階と、複数の動画像情報とカーソル画像情報とを合成して表示画像を生成しこれを前記表示装置上に出力する段階と、複数の動画像情報の表示位置とカーソル位置情報との間の距離を算出し距離情報を生成する段階と、距離情報に基づいて複数の動画像情報に対応する音声データの音量を決定し音声出力装置に出力する段階とを含むマルチチャネル情報処理方法をコンピュータに実行させるためのプログラムである。

本発明の請求項 17 に係るプログラムは、複数の動画像情報を同時に表示装置に表示する際のマルチチャネル情報処理方法のプログラムであって、複数の動画像情報の表示装置上での表示位置を決定する段階と、決定した表示位置に基づいて複数の動画像情報を出力する段階と、入力デバイスから入力されるカーソル指示情報を受け付ける段階と、カーソル指示情報に基づいてカーソルを表示するカーソル位置情報を算出する段階と、カーソル位置情報に基づいてカーソル画像情報を生成する段階と、複数の動画像情報とカーソル画像情報とを合成して表示画像を生成しこれを表示装置上に出力する段階と、複数の動画像情報の表示位置とカーソル位置情報との間の距離を算出し距離情報を生成する段階と、複数の動画像情報について生成された距離情報のうち距離の値が最も小さい動画像情報に対応する音声データの音量を予め定められた音量  $V_a$  に設定し、他の動画像情報に対応する音声データの音量を音量  $V_a$  よりも小さな予め定められた音量  $V_b$  に設定し音声出力装置に出力する段階とを含むマルチチャネル情報処理方法をコンピュータに実行させるためのプログラムである。

本発明の請求項 18 に係るプログラムは、複数の動画像情報を同時に表示装置

に表示する際のマルチチャネル情報処理方法のプログラムであって、複数の動画  
像情報の表示装置上での表示位置を決定する段階と、決定した表示位置に基づい  
て複数の動画像情報を出力する段階と、入力デバイスから入力されるカーソル指  
示情報を受け付ける段階と、カーソル指示情報に基づいてカーソルを表示するカ  
ーソル位置情報を算出する段階と、カーソル位置情報に基づいてカーソル画像情  
報を生成する段階と、複数の動画像情報とカーソル画像情報とを合成して表示画  
像を生成しこれを表示装置上に出力する段階と、複数の動画像情報の表示位置と  
カーソル位置情報との間の距離を算出し距離情報を生成する段階と、複数の動画  
像情報に対応する音声データの音量を、各距離情報中の距離の値に反比例するよ  
うに設定し音声出力装置に出力する段階とを含むマルチチャネル情報処理方法を  
コンピュータに実行させるためのプログラムである。

本発明の請求項 19 に係るプログラムは、複数の動画像情報を同時に表示装置  
に表示する際のマルチチャネル情報処理方法のプログラムであって、複数の動画  
像情報の表示装置上での表示位置を決定する段階と、決定した表示位置に基づい  
て複数の動画像情報を出力する段階と、入力デバイスから入力されるカーソル指  
示情報を受け付ける段階と、カーソル指示情報に基づいてカーソルを表示するカ  
ーソル位置情報を算出する段階と、カーソル位置情報に基づいてカーソル画像情  
報を生成する段階と、複数の動画像情報とカーソル画像情報とを合成して表示画  
像を生成しこれを表示装置上に出力する段階と、複数の動画像情報の表示位置と  
カーソル位置情報との間の距離を算出し距離情報を生成する段階と、カーソルの  
表示位置から見た各動画像情報の表示位置の方向に関する方向情報を生成する段  
階と、複数の動画像情報に対応する音声データが、音声出力装置の音像空間にお  
ける距離情報および方向情報に応じた音像位置に定位するように、音声出力装置  
に出力を行う段階とを含むマルチチャネル情報処理方法をコンピュータに実行さ  
せるためのプログラムである。

本発明の請求項 20 に係るプログラムは、複数の動画像情報を同時に表示装置  
に表示する際のマルチチャネル情報処理方法のプログラムであって、複数の動画

像情報の表示装置上での表示位置を決定する段階と、決定した表示位置に基づいて複数の動画像情報を出力する段階と、入力デバイスから入力されるカーソル指示情報を受け付ける段階と、カーソル指示情報に基づいてカーソルを表示するカーソル位置情報を算出する段階と、カーソル位置情報に基づいてカーソル画像情報を生成する段階と、複数の動画像情報に対応する音声データに含まれる言葉を音声認識する段階と、音声認識した言葉を文字情報に変換してこれを出力する段階と、複数の動画像情報、カーソル画像情報および文字情報とを合成して表示画像を生成しこれを表示装置上に出力する段階とを備えるマルチチャンネル情報処理方法をコンピュータに実行させるためのプログラムである。

10

本発明の請求項 2 1 に係るプログラムは、複数の動画像情報を同時に表示装置に表示する際のマルチチャンネル情報処理方法のプログラムであって、複数の動画像情報の表示装置上での表示位置を決定する段階と、決定した表示位置に基づいて複数の動画像情報を出力する段階と、入力デバイスから入力されるカーソル指示情報を受け付ける段階と、カーソル指示情報に基づいてカーソルを表示するカーソル位置情報を算出する段階と、カーソル位置情報に基づいてカーソル画像情報を生成する段階と、複数の動画像情報の表示位置とカーソル位置情報との間の距離を算出し距離情報を生成する段階と、距離情報に基づいて複数の動画像情報のうちから特定の動画像情報を選択し、この動画像情報に対応する音声データを音声出力装置に出力する段階と、音声出力装置に出力する音声データに含まれる言葉を音声認識する段階と、音声認識した言葉を文字情報に変換してこれを出力する段階と、複数の動画像情報、カーソル画像情報および文字情報とを合成して表示画像を生成しこれを表示装置上に出力する段階とを備えるマルチチャンネル情報処理方法をコンピュータに実行させるためのプログラムである。

25

本発明の請求項 2 2 に係るプログラムは、複数の動画像情報を同時に表示装置に表示する際のマルチチャンネル情報処理方法のプログラムであって、複数の動画像情報の表示装置上での表示位置を決定する段階と、決定した表示位置に基づいて複数の動画像情報を出力する段階と、入力デバイスから入力されるカーソル指



- 示情報を受け付ける段階と、カーソル指示情報に基づいてカーソルを表示するカーソル位置情報を算出する段階と、カーソル位置情報に基づいてカーソル画像情報を生成する段階と、複数の動画像情報の表示位置とカーソル位置情報との間の距離を算出し距離情報を生成する段階と、距離情報に基づいて複数の動画像情報のうちから特定の動画像情報を選択し、この動画像情報に対応する音声データを音声出力装置に出力する段階と、音声出力装置に出力する音声データに含まれる言葉を音声認識する段階と、インターネットに接続する段階と、音声認識した言葉をキーワードとしてインターネット上に存在する関連ホームページを検索する段階と、複数の動画像情報、カーソル画像情報および検索された関連ホームページを合成して表示画像を生成しこれを表示装置上に出力する段階とを備えるマルチチャンネル情報処理方法をコンピュータに実行させるためのプログラムである。

- 本発明の請求項 2 3 に係るプログラムは、複数の動画像情報を同時に表示装置に表示する際のマルチチャンネル情報処理方法のプログラムであって、複数の動画像情報の表示装置上での表示位置を決定する段階と、決定した表示位置に基づいて複数の動画像情報を出力する段階と、表示装置に表示される複数の動画像情報から所定のアルゴリズムに基づいて特定の動画像情報を選択する段階と、選択した動画像情報に対応する音声データを音声出力装置に出力する段階とを備えるマルチチャンネル情報処理方法をコンピュータに実行させるためのプログラムである。

本発明の請求項 2 4 に係る記憶媒体は、請求項 1 6 ～ 2 3 のプログラムを記憶したコンピュータが読み取り可能な記憶媒体である。

(図面の簡単な説明)

- 図 1 は、本発明の概略構成を示すブロック図である。  
図 2 は、動画像位置情報テーブルの一例を示す説明図である。  
図 3 は、画面構成の一例を示す説明図である。  
図 4 は、第 1 実施形態の制御ブロック図である。  
図 5 は、第 1 実施形態の制御フローチャートである。

図 6 は、動画像一情報テーブルの他の例を示す説明図である。

図 7 は、第 1 実施形態の変形例の制御フローチャートである。

図 8 は、第 2 実施形態の制御ブロック図である。

図 9 は、第 2 実施形態の制御フローチャートである。

5 図 10 は、第 3 実施形態の制御ブロック図である。

図 11 は、第 3 実施形態の制御フローチャートである。

図 12 は、第 3 実施形態の変形例の制御ブロック図である。

図 13 は、第 4 実施形態の制御ブロック図である。

図 14 は、第 4 実施形態の制御フローチャートである。

10

(発明を実施するための最良の形態)

〔基本構成〕

本発明の 1 実施形態が採用される情報処理装置のハード構成を図 1 に示す。

15 情報処理装置 10 は、通常のパーソナルコンピュータで構成することが可能であり、PC 本体 1、CRT や液晶表示装置などで構成されるディスプレイ 2、スピーカ 3、マウスやトラックボール、タッチパッド、キーボードなどの入力デバイス 4、インターネット接続が可能なモデム 5 などを備えている。モデム 5 は、たとえば公衆回線を通じてインターネット接続を行う ADSL モデムなどが採用される。

20 PC 本体 1 は、モデム 5 および公衆回線を介して ISP (Internet Services Provider) 20 に接続を行い、ISP 20 に用意されているストリーミングデータから複数の動画像情報をダウンロードすることが可能な構成となっている。ダウンロードされた複数の動画像情報は、ディスプレイ 2 上の予めプログラムで設定されている表示位置に表示される。もちろん、プログラム上に用意されたツール

25 を用いてユーザが表示位置を設定するように構成することも可能である。

動画像情報の表示位置に関する位置情報テーブルの一例を図 2 に示す。

この位置情報テーブルは、チャンネル番号、中心の位置、左右の幅、上下の高さなどの項目を備えており、各チャンネル番号  $CH_i$  ( $i=1\sim n$ ) の動画像情報の中心の座標位置 ( $CX_i$ ,  $CY_i$ )、左右の幅  $W_i$ 、上下の高さ  $H_i$  がそれぞれ格

納される。この位置情報テーブルは、表示されるチャンネル数 $n$ 毎に設定されており、PC本体1で処理可能な最大チャンネル数またはディスプレイ2上に表示可能な最大チャンネル数まで用意されている。

- 5      このような位置情報テーブルで設定された表示位置に基づいて動画像情報の表示を行う場合の表示画面の一例を図3に示す。ここでは、表示チャンネル数が3の場合を示しており、それぞれ中心の座標位置( $CX_i$ ,  $CY_i$ )、左右の幅 $W_i$ 、上下の高さ $H_i$ によって、各動画像情報の表示位置を決定することができる。

#### 〔第1実施形態〕

- 10      本発明の第1実施形態の制御ブロック図を図4に示す。

この情報処理装置100は、動画像源200から複数の動画像情報を取得する動画像表示制御手段101を備えている。動画像情報表示制御手段101は、たとえば、図2に示されるような位置情報テーブルにしたがって各動画像情報の表示位置を決定し、この動画像位置情報と動画像情報とを出力する。

- 15      カーソル位置決定手段102は、マウス、トラックボール、タッチパッド、キーボード、その他のポインティングデバイスで構成される入力デバイス140により入力されるカーソル指示情報を受け付け、このカーソル指示情報と現在のカーソルの表示位置とに基づいて次のカーソル位置情報を算出しこれを出力する。

- 20      カーソル位置決定手段102から出力されるカーソル位置情報はカーソル画像生成手段103に入力される。カーソル画像生成手段103は、カーソル位置情報に基づいてカーソル画像情報を生成しこれを出力する。

- 25      動画像表示制御手段101から出力される動画像情報と動画像位置情報およびカーソル画像生成手段103から出力されるカーソル画像情報は、表示画像生成手段104に入力される。表示画像生成手段104は、VRAMなどの画像出力用バッファを備え、複数の動画像情報をその位置情報に基づいて配置するとともに、カーソル画像を合成して表示装置120に出力する。

動画像表示制御手段101から出力される動画像位置情報とカーソル位置決定手段102から出力されるカーソル位置情報は距離情報生成手段105に入力される。距離情報生成手段105は、動画像位置情報とカーソル位置情報に基づい

て各動画像情報の表示位置とカーソルの表示位置との距離を算出する。この各動画像情報の表示位置とカーソルの表示位置との距離は、動画像情報の中心位置とカーソルの中心位置との間の距離とすることができる。また、カーソルが動画像情報のウィンドウ内に位置する場合には、その距離を” 0 ” に設定するように構成

5 成できる。

距離情報生成手段 105 で生成された距離情報は動画像選択手段 106 に入力される。動画像選択手段 106 では、距離情報生成手段 105 で生成された距離情報に基づいて、どの動画像情報に対応する音声データを出力するかを決定する。たとえば、各動画像情報の距離情報のうち最も距離の値が小さい動画像情報を

10 選択するように構成できる。

動画像選択手段 106 は、選択された動画像情報に基づいて音声出力選択情報を音声出力選択手段 107 に出力する。音声出力選択手段 107 は、動画像源 200 からの複数の動画像情報のうち、音声出力選択情報で設定される動画像情報に対応する音声データを選択的に取得し、これを音声出力装置 130 に出力する

15 。

この構成における動作を図 5 のフローチャートに基づいて説明する。

ステップ S 11 では、変数  $i$  の値を初期化する。

ステップ S 12 では、カーソル位置情報の読み込みを行う。

20 ステップ S 13 では、カーソルの表示位置がチャンネル番号  $CH_i$  の動画像情報の表示枠内に入っているか否かを判別する。動画像情報の位置情報テーブルとして、図 6 に示すような、チャンネル番号、中心の位置、表示幅、表示高さの各項目に加えて、近傍幅、近傍高さの項目を備えるものを用いることができる。カーソルの表示位置が、チャンネル番号  $CH_i$  の動画像情報の中央位置座標 ( $CX_i$ ,  $CY_i$ ) および近傍幅  $NW_i$ 、近傍高さ  $NH_i$  で規定される範囲内にある場合には、この動画像情報の表示枠に入っていると判断する。この場合、動画像情報を表示しているウィンドウの周囲に所定の距離範囲で表示枠を設定し、この表示枠内にカーソルが入った場合にその動画像情報が選択されたと判断するように構成できる。カーソルの表示位置がチャンネル番号  $CH_i$  の動画像情報の表示枠内に入っていると判断した場合にはステップ S 16 に移行する。

ステップS 1 6では、チャンネル番号CH iの動画像情報に対応する音声データを音声出力装置1 3 0に出力する。

ステップS 1 3において、カーソルの表示位置がチャンネル番号CH iの動画像情報の表示枠内に入っていないと判断した場合には、ステップS 1 4に移行する

5 。

ステップS 1 4では、変数iの値が表示されている動画像情報数nに到達したか否かを判別する。変数i = 表示チャンネル数nである場合にはステップS 1 1に移行し、そうでない場合にはステップS 1 5に移行する。

10 ステップS 1 5では変数iの値をインクリメントしステップS 1 3に移行する。

このように構成した場合、たとえばマウスカーソルの表示位置が動画像情報のウィンドウ上にない場合であっても、ウィンドウの近傍にあることを認識して、その動画像情報に対応する音声データを出力するように構成できる。ユーザは、マウスやトラックボール、その他のポインティングデバイスを用いてカーソルを移動させ、希望する動画像情報のウィンドウにカーソルを近づけることで、その音声データを聞くことができ、動画像情報の内容を容易に認識することが可能となる。

#### 〈変形例1〉

20 カーソルの表示位置に最も近い動画像情報を選択して音声データを出力するように構成することも可能である。この場合のフローチャートを図7に示す。

ステップS 2 1では、変数iの値を初期化する。

ステップS 2 2では、カーソル位置情報の読み込みを行う。

25 ステップS 2 3では、カーソルの表示位置とチャンネル番号CH iの動画像情報の表示位置との距離d iを算出する。この距離d iは、動画像情報を表示しているウィンドウの中心座標とカーソルの中心座標との間の距離とすることも可能であり、また動画像情報を表示しているウィンドウ枠とカーソルの中心座標との最短距離とすることも可能である。

ステップS 2 4では、変数iの値が表示されている動画像情報数nに到達した

か否かを判別する。変数  $i$  = 表示チャンネル数  $n$  である場合にはステップ S 2 6 に移行し、そうでない場合にはステップ S 2 5 に移行する。

ステップ S 2 5 では、変数  $i$  の値をインクリメントしステップ S 2 3 に移行する。

- 5     ステップ S 2 6 では、動画像情報の表示位置とカーソルの表示位置との間の距離  $d_i$  のうち最小値を判定し、これに対応するチャンネル  $CH_i$  の動画像情報を選択画像に決定する。

ステップ S 2 7 では、選択された動画像情報に対応する音声データの出力を行う。

- 10    このように構成した場合、動画像情報を表示するウィンドウの近傍に表示枠を設けるためのテーブルを用意する必要がなく、カーソルの表示位置に最も近い動画像情報に対応する音声データを出力するように構成できる。

#### 〈変形例 2〉

- 15    音声データの出力を行っている動画像情報について、その表示ウィンドウの枠の色を変更したり、選択されていない動画像情報の色調を変更するなどにより、選択されたチャンネルを認識できるように構成することも可能である。

#### 〈変形例 3〉

- 20    動画像情報の表示位置とカーソルの表示位置との間の距離  $d_i$  に反比例するように、各動画像情報に対応する音声データを増幅し、各音声データを合成して  $\Sigma$  出力するように構成することも可能である。

カーソルの表示位置の中心座標を  $(C_uX, C_uY)$ 、チャンネル  $CH_i$  の動画像情報の中心座標  $(CX_i, CY_i)$  としこれらの座標間の距離  $d_i$  とすると、

25    
$$d_i = ((C_uX - CX_i)^2 + (C_uY - CY_i)^2)^{1/2}$$

で計算することができる。この距離  $d_i$  を用いてその動画像情報に対応する音声データの増幅率  $CHig[dB]$  を、

$$CHig[dB] = a \cdot d_i \quad (\text{ただし、} a < 0 \text{ の定数})$$

で設定する。

- このことにより、カーソル表示位置に近い動画像情報の音声データは大きな音量で出力され、カーソル表示位置から遠い動画像情報の音声データは小さい音量で出力されることとなる。したがって、マウスなどのポインティングデバイスを用いてカーソルを近づけた動画像情報の音声データの音量が大きくなり、その動画像情報の内容把握を容易にするとともに、他の動画像情報についても音量が小さいながらも音声データの出力を行っているため、内容把握の助けとすることが可能である。

#### 〈変形例 4〉

- 10      カーソルの表示位置から見た動画像情報の表示位置の方向に関する方向情報を算出し、音声出力装置 130 が生成する音像空間内に方向情報と距離情報に基づいて動画像情報に対応する音声データの音像を定位するように構成することも可能である。

- 15      前述したように、カーソルの表示位置および動画像情報の表示位置は、それぞれ表示装置上の座標で表すことが可能であり、この座標を用いてカーソルから見た動画像情報の表示位置の方向情報を算出することが可能である。このような方向情報と距離情報  $d_i$  を用いて、音声出力装置 130 が生成する音像空間内に音声データの音像を定位することが可能となる。ただし、この場合、音声出力装置 130 は 2 以上のスピーカを備え、各スピーカから出力される音声による音像空間が生成されるように各スピーカからの出力が制御される必要がある。

- 20      たとえば、カーソルの中心座標  $(C_u X, C_u Y)$ 、チャンネル  $CH_i$  の中心座標  $(C X_i, C Y_i)$  とし、X 軸方向の距離  $C X_{id}$  を、

$$C X_{id} = C_u X - C X_i$$

- 25      として算出する。この X 軸方向の距離  $C X_{id}$  を用いて、左右のスピーカに出力する音量を決定して出力することにより、音像の定位を行うことが可能となる。

このような構成とすることにより、表示装置 120 の動画像情報の表示位置と音声出力装置 130 から出力される音声データの位置との間に整合性を持たせることができ、動画像情報の内容把握を促すことが容易になる。

## 〔第2実施形態〕

本発明の第2実施形態の制御ブロック図を図8に示す。

この情報処理装置150は、動画像源200から複数の動画像情報を取得する動画像表示制御手段101を備えている。動画像情報表示制御手段101は、第1実施形態と同様に、予め設定された位置情報テーブルにしたがって各動画像情報の表示位置を決定し、この動画像位置情報と動画像情報とを出力する。

表示画像生成手段104は、第1実施形態と同様に、VRAMなどの画像出力用バッファを備え、複数の動画像情報をその位置情報に基づいて配置し画像出力を行う。

動画像表示制御手段101からの動画像位置情報は動画像選択手段108に入力される。動画像選択手段108は、複数のチャンネルの動画像情報を所定の周期で選択するようなアルゴリズムを有している。たとえば、ひとつのチャンネル番号を選択してこの動画像情報に対応する音声データを選択するための音声選択情報を出し、一定時間経過する毎にチャンネル番号順に選択するチャンネル番号を切り換えていくように構成できる。

動画像選択手段108から出力される音声出力選択情報は音声出力選択手段107に入力される。音声出力選択手段107は、音声出力選択情報に基づいて選択されているチャンネルの動画像情報に対応する音声データを音声出力装置130に出力する。

表示されている動画像情報を順に選択し、選択された動画像情報に対応する音声データを一定時間出力するような構成のアルゴリズムを図9のフローチャートに示す。

ステップS31では、変数*i*の値を初期化する。

ステップS32では、タイマを初期化してこのタイマによる経過時間のカウンタを開始する。このタイマは、選択されたチャンネルの動画像情報に対応する音声データを出力する時間を計時するものである。

ステップS33では、チャンネルCH*i*の画像情報に対応する音声データを音声出力装置130に出力する。

ステップS34では、タイマのカウント値が所定値に達したか否かを判別する



。タイマのカウント値が予め設定されている所定値に到達したと判断した場合にはステップS 3 5に移行する。

ステップS 3 5では、変数*i*の値をインクリメントし、次のチャンネル番号の動画像情報を選択する。

- 5     ステップS 3 6では、変数*i*の値が表示を行っているチャンネル数*n*を超えたか否かを判別する。変数*i*の値がチャンネル数*n*を超えていないと判断した場合にはステップS 3 2に移行し、変数*i*の値がチャンネル数*n*を超えたと判断した場合にはステップS 3 1に移行する。

- 10    動画像選択手段1 0 8がこのようなアルゴリズムにしたがって動画像情報を選択し、選択した動画像情報に対応する音声データを出力するように構成することで、周期的に選択する動画像情報を切り換えてそれに対応する音声データを一定時間出力することができる。ユーザは、各動画像情報をポインティングデバイスなどで指定する作業なしで、各動画像情報に対応する音声データを周期的に認識することが可能であり、各動画像情報の内容を把握することが容易となる。

15

### 〔第3実施形態〕

本発明の第3実施形態の制御ブロック図を図1 0に示す。

- 20    この情報処理装置1 6 0は、動画像源2 0 0から複数の動画像情報を取得する動画像表示制御手段1 0 1を備えている。動画像情報表示制御手段1 0 1は、第1実施形態と同様にして各動画像情報の表示位置を決定し、この動画像位置情報と動画像情報とを出力する。

- 25    カーソル位置決定手段1 0 2は、マウス、トラックボール、タッチパッド、キーボード、その他のポインティングデバイスで構成される入力デバイス1 4 0により入力されるカーソル指示情報を受け付け、このカーソル指示情報と現在のカーソルの表示位置とに基づいて次のカーソル位置情報を算出しこれを出力する。

カーソル位置決定手段1 0 2から出力されるカーソル位置情報はカーソル画像生成手段1 0 3に入力される。カーソル画像生成手段1 0 3は、カーソル位置情報に基づいてカーソル画像情報を生成しこれを出力する。

動画像表示制御手段1 0 1から出力される動画像位置情報とカーソル位置決定

手段102から出力されるカーソル位置情報は距離情報生成手段105に入力される。距離情報生成手段105は、動画像位置情報とカーソル位置情報に基づいて各動画像情報の表示位置とカーソルの表示位置との距離を算出する。この各動画像情報の表示位置とカーソルの表示位置との距離は、動画像情報の中心位置とカーソルの中心位置との間の距離とすることができる。また、カーソルが動画像情報のウィンドウ内に位置する場合には、その距離を”0”に設定するように構成できる。

距離情報生成手段105で生成された距離情報は動画像選択手段106に入力される。動画像選択手段106では、距離情報生成手段105で生成された距離情報に基づいて、どの動画像情報に対応する音声データを出力するかを決定する。たとえば、各動画像情報の距離情報のうち最も距離の値が小さい動画像情報を選択するように構成できる。

動画像選択手段106は、選択された動画像情報に基づいて音声出力選択情報を音声出力選択手段107に出力する。音声出力選択手段107は、動画像源200からの複数の動画像情報のうち、音声出力選択情報で設定される動画像情報に対応する音声データを選択的に取得し、これを音声出力装置130に出力する。

音声出力選択手段107から出力される音声データは、音声認識手段109に入力される。音声認識手段109は、入力される音声データから意味のある言葉を認識してこれを抽出し文字情報に変換するものである。

音声認識手段109で変換された文字情報は認識文字表示手段110に入力される。認識文字表示手段110は、音声認識手段109から入力される文字情報に基づいて表示装置120上に表示する文字データを生成し表示画像生成手段104に出力する。

動画像表示制御手段101から出力される動画像情報と動画像位置情報、カーソル画像生成手段103から出力されるカーソル画像情報および認識文字表示手段110から出力される文字データは、表示画像生成手段104に入力される。表示画像生成手段104は、VRAMなどの画像出力用バッファを備え、複数の動画像情報をその位置情報に基づいて配置し、カーソル画像および文字データを

合成して表示装置 120 に出力する。

ここで、認識文字表示手段 110 から出力される文字データは、対応する動画像情報の近傍に表示するように構成することが可能であり、どの動画像情報から得た文字データであるかを示す情報を表示装置 120 上に明示するような構成と

5    することも可能である。

また、選択されたチャンネルの動画像情報に対応する音声データから取得した文字データを表示し、音声出力装置 130 から音声データの出力を行わない構成とすることも可能である。

10    この場合の動作について図 11 に示すフローチャートに基づいて説明する。

ステップ S41 では、変数  $i$  の値を初期化する。

ステップ S12 では、カーソル位置情報の読み込みを行う。

ステップ S13 では、カーソルの表示位置がチャンネル番号  $CH_i$  の動画像情報の表示枠内に入っているか否かを判別する。この場合、第 1 実施形態と同様に、

15    動画像情報を表示しているウィンドウの周囲に所定の距離範囲で表示枠を設定し、この表示枠内にカーソルが入った場合にその動画像情報が選択されたと判断するように構成できる。カーソルの表示位置がチャンネル番号  $CH_i$  の動画像情報の表示枠内に入っていると判断した場合にはステップ S44 に移行する。

20    ステップ S44 では、チャンネル番号  $CH_i$  の動画像情報に対応する音声データを音声出力装置 130 に出力する。

ステップ S45 では、チャンネル番号  $CH_i$  の動画像情報に対応する音声データから意味のある言葉を認識し文字データに変換する。

ステップ S46 では、認識した文字データを表示画像生成手段 104 に送出し、表示装置 120 上に表示する。

25    ステップ S43 において、カーソルの表示位置がチャンネル番号  $CH_i$  の動画像情報の表示枠内に入っていないと判断した場合には、ステップ S47 に移行する。

ステップ S47 では、変数  $i$  の値が表示されている動画像情報数  $n$  に到達したか否かを判別する。変数  $i$  = 表示チャンネル数  $n$  である場合にはステップ S41 に

移行し、そうでない場合にはステップ S 4 8 に移行する。

ステップ S 4 8 では変数  $i$  の値をインクリメントしステップ S 4 3 に移行する

。

- 5      このように構成した場合には、選択されたチャンネルの動画像情報に対応する音声データに含まれる意味のある言葉を視覚的に認識することが可能となり、動画像情報の内容把握を容易にすることができる。

#### 〈変形例〉

第 3 実施形態の変形例のブロック図を図 1 2 に示す。

- 10      この情報処理装置 1 7 0 では、音声認識手段 1 0 9 および認識文字表示手段 1 1 0 が表示するチャンネル数だけ設けられており、それぞれ各チャンネルの動画像情報に対応する音声データから意味のある言葉を抽出して文字データに変換しこれを表示画像生成手段 1 0 4 に入力し、表示装置 1 2 0 上に表示する構成となっている。

- 15      各文字データはそれぞれ対応する動画像情報の表示位置の近傍に表示されるように構成することが可能であり、文字データのみをリスト表示するように構成することも可能である。

- 20      この場合には、各動画像情報に対応する音声データから抽出された言葉を表示装置 1 2 0 上にそれぞれ表示することが可能となり、各動画像情報の内容把握を容易にすることができる。

#### 〔第 4 実施形態〕

本発明の第 4 実施形態の制御ブロック図を図 1 3 に示す。

- 25      この情報処理装置 1 8 0 は、動画像源 2 0 0 から複数の動画像情報を取得する動画像表示制御手段 1 0 1 を備えている。動画像情報表示制御手段 1 0 1 は、第 1 実施形態と同様にして各動画像情報の表示位置を決定し、この動画像位置情報と動画像情報とを出力する。

カーソル位置決定手段 1 0 2 は、マウス、トラックボール、タッチパッド、キーボード、その他のポインティングデバイスで構成される入力デバイス 1 4 0 に

より入力されるカーソル指示情報を受け付け、このカーソル指示情報と現在のカーソルの表示位置とに基づいて次のカーソル位置情報を算出しこれを出力する。

カーソル位置決定手段102から出力されるカーソル位置情報はカーソル画像生成手段103に入力される。カーソル画像生成手段103は、カーソル位置情報に基づいてカーソル画像情報を生成しこれを出力する。

動画像表示制御手段101から出力される動画像位置情報とカーソル位置決定手段102から出力されるカーソル位置情報は距離情報生成手段105に入力される。距離情報生成手段105は、動画像位置情報とカーソル位置情報に基づいて各動画像情報の表示位置とカーソルの表示位置との距離を算出する。この各動画像情報の表示位置とカーソルの表示位置との距離は、動画像情報の中心位置とカーソルの中心位置との間の距離とすることができる。また、カーソルが動画像情報のウィンドウ内に位置する場合には、その距離を"0"に設定するように構成できる。

距離情報生成手段105で生成された距離情報は動画像選択手段106に入力される。動画像選択手段106では、距離情報生成手段105で生成された距離情報に基づいて、どの動画像情報に対応する音声データを出力するかを決定する。たとえば、各動画像情報の距離情報のうち最も距離の値が小さい動画像情報を選択するように構成できる。

動画像選択手段106は、選択された動画像情報に基づいて音声出力選択情報を音声出力選択手段107に出力する。音声出力選択手段107は、動画像源200からの複数の動画像情報のうち、音声出力選択情報で設定される動画像情報に対応する音声データを選択的に取得し、これを音声出力装置130に出力する。

音声出力選択手段107から出力される音声データは、音声認識手段109に入力される。音声認識手段109は、入力される音声データから意味のある言葉を認識してこれを抽出し文字情報に変換するものである。

音声認識手段109で変換された文字情報は認識文字表示手段110に入力される。認識文字表示手段110は、音声認識手段109から入力される文字情報に基づいて表示装置120上に表示する文字データを生成し表示画像生成手段1

04に出力する。

音声認識手段109から出力される文字情報はホームページ検索手段111に入力される。ホームページ検索手段111はインターネット接続手段113を介してインターネットにアクセスを行い、音声認識手段109から取得した文字情報

- 5 報をキーワードとしてインターネット上で検索可能なホームページを検索する。ホームページの検索には、YAHOO、goo、googleなどの検索サイトに接続して検索結果を受信するように構成することが可能である。検索結果はインターネット接続手段113を介してホームページ検索手段111で受信され、ホームページ表示手段112に送出される。

- 10 ホームページ表示手段112は、検索結果から得られたホームページのURLにアクセスを行い、ホームページの情報を取得して表示画像生成手段104に出力する。ホームページ表示手段112は、Internet ExplorerやNetscape Navigatorなどのウェブブラウザで構成することが可能である。また、ホームページ情報の取得はインターネット接続手段113を介してインターネットにアクセスを行
- 15 うことで可能となる。

- 動画像表示制御手段101から出力される動画像情報と動画像位置情報、カーソル画像生成手段103から出力されるカーソル画像情報、認識文字表示手段110から出力される文字データおよびホームページ表示手段112から出力されるホームページ情報は、表示画像生成手段104に入力される。表示画像生成手段104は、VRAMなどの画像出力用バッファを備え、複数の動画像情報をその位置情報に基づいて配置し、カーソル画像、文字データおよびホームページ情報を合成して表示装置120に出力する。
- 20

この場合の動作について、図14に示すフローチャートに基づいて説明する。

- 25 ステップS51では、変数iの値を初期化する。

ステップS52では、カーソル位置情報の読み込みを行う。

ステップS53では、カーソルの表示位置がチャンネル番号CHiの動画像情報の表示枠内に入っているか否かを判別する。この場合、第1実施形態と同様に、動画像情報を表示しているウィンドウの周囲に所定の距離範囲で表示枠を設定し

、この表示枠内にカーソルが入った場合にその動画像情報が選択されたと判断するように構成できる。カーソルの表示位置がチャンネル番号CH*i*の動画像情報の表示枠内に入っていると判断した場合にはステップS54に移行する。

5     ステップS54では、チャンネル番号CH*i*の動画像情報に対応する音声データを音声出力装置130に出力する。

ステップS55では、チャンネル番号CH*i*の動画像情報に対応する音声データから意味のある言葉を認識し文字データに変換する。

ステップS56では、認識した文字データを表示画像生成手段104に送出し、表示装置120上に表示する。

10    ステップS57では、音声認識により得られた文字データからホームページ検索を行う。

ステップS58では、ホームページ検索を行った検索結果に基づいて、ホームページ情報を取得し、これを表示装置120上に表示する。

15    ステップS53において、カーソルの表示位置がチャンネル番号CH*i*の動画像情報の表示枠内に入っていないと判断した場合には、ステップS59に移行する。

ステップS59では、変数*i*の値が表示されている動画像情報数*n*に到達したか否かを判別する。変数*i*=表示チャンネル数*n*である場合にはステップS51に移行し、そうでない場合にはステップS60に移行する。

20    ステップS60では変数*i*の値をインクリメントしステップS53に移行する。

25    このように構成した場合には、動画像情報に対応する音声データから得られる言葉から関連ホームページの情報を取得してこれを表示装置上に表示させることが可能となる。このことにより、各動画像情報の関連情報を容易に取得することができ、種々のデータを簡単に取得することができる。

#### (産業上の利用可能性)

本発明によれば、表示装置に表示される複数の動画像情報から特定の動画像情報に対応する音声データを容易に認識することができ、その動画像情報の内容の

把握が容易になる。



## 請 求 の 範 囲

1.

5 複数の動画像情報を同時に表示装置に表示するマルチチャネル情報処理装置であって、

前記複数の動画像情報を取得し、前記複数の動画像情報の表示装置上での表示位置に関する動画像位置情報を決定して前記複数の動画像情報を前記動画像位置情報に基づいて出力する動画像情報制御手段と、

10 入力デバイスを介して入力されるカーソル指示情報に基づいてカーソル位置情報を算出し前記カーソル位置情報に基づいてカーソル画像情報を生成してこれを出力するカーソル位置制御手段と、

前記動画像表示制御手段が出力する複数の動画像情報と、前記カーソル位置制御手段が出力するカーソル画像情報とを合成して前記表示装置上に表示する表示画像生成手段と、

15 前記複数の動画像情報に対応する動画像位置情報と、前記カーソル位置制御手段で算出されたカーソル位置情報とに基づいて、各動画像情報の表示位置とカーソルの表示位置との距離を算出し距離情報を生成する距離情報生成手段と、

前記距離情報生成手段で生成された距離情報に基づいて前記複数の動画像情報に対応する音声データの音量を決定して音声出力装置に出力する音声出力制御手段と、

20 を備えるマルチチャネル情報処理装置。

2.

25 前記音声出力制御手段は、前記距離情報生成手段で生成された距離情報のうち距離の値が最も小さい動画像情報に対応する音声データの音量を予め定められた音量  $V_a$  に設定し、他の動画像情報に対応する音声データの音量を前記音量  $V_a$  よりも小さな音量  $V_b$  に設定することを特徴とする、請求項 1 に記載のマルチチャネル情報処理装置。

3.

前記音声出力制御手段は、前記複数の動画像情報に対応する音声データの音量を、前記距離情報生成手段で生成された距離情報中の距離の値に反比例するように設定することを特徴とする、請求項1に記載のマルチチャネル情報処理装置。

5

4.

前記距離情報生成手段で生成された距離情報は、カーソルの表示位置から見た各動画像情報の表示位置の方向に関する方向情報を含み、前記音声出力制御手段は前記音声出力装置が形成する音像空間に前記複数の動画像情報に対応する音声データが定位するように前記距離情報に基づいて前記音声出力装置に出力を行うことを特徴とする、請求項1に記載のマルチチャネル情報処理装置。

10

5.

前記複数の動画像情報に対応する音声データに含まれる言葉を認識する音声データ認識手段と、

15

前記音声データ認識手段で認識した言葉を文字情報に変換して前記表示装置に表示する文字情報表示手段と、  
をさらに備える請求項1に記載のマルチチャネル情報処理装置。

20

6.

インターネット接続手段と、

前記音声データ認識手段で認識した言葉をキーワードとしてインターネット上に存在する関連ホームページを検索するホームページ検索手段と、

25

前記ホームページ検索手段で検索されたホームページを前記表示装置上に表示するホームページ表示手段と、  
を備える、請求項5に記載のマルチチャネル情報処理装置。

7.

前記表示装置に表示される複数の動画像情報から所定のアルゴリズムに基づい

て特定の動画像情報を選択する動画像選択手段をさらに備え、前記音声出力制御手段は前記動画像選択手段が選択した動画像情報に対応する音声データを音声出力装置に出力することを特徴とする、請求項 1 に記載のマルチチャンネル情報処理装置。

5

8.

前記動画像選択手段は、所定時間毎に選択する動画像情報を切り換えることを特徴とする、請求項 7 に記載のマルチチャンネル情報処理装置。

10

9.

複数の動画像情報を同時に表示装置に表示する際のマルチチャンネル情報処理方法であって、

前記表示装置上に表示を行う動画像情報の表示装置上での表示位置を決定する段階と、

15

決定した表示位置に基づいて前記複数の動画像情報を出力する段階と、

前記入力デバイスから入力されるカーソル指示情報を受け付ける段階と、

前記カーソル指示情報に基づいてカーソルを表示するカーソル位置情報を算出する段階と、

前記カーソル位置情報に基づいてカーソル画像情報を生成する段階と、

20

前記複数の動画像情報と前記カーソル画像情報とを合成して表示画像を生成しこれを前記表示装置上に出力する段階と、

前記複数の動画像情報の表示位置と前記カーソル位置情報との間の距離を算出し距離情報を生成する段階と、

前記距離情報に基づいて複数の動画像情報に対応する音声データの音量を決定

25

し音声出力装置に出力する段階と、

を含むマルチチャンネル情報処理方法。

10.

複数の動画像情報について生成された距離情報のうち距離の値が最も小さい動

画像情報に対応する音声データの音量を予め定められた音量 $V_a$ に設定し、他の動画像情報に対応する音声データの音量を前記音量 $V_a$ よりも小さな予め定められた音量 $V_b$ に設定することを特徴とする、請求項9に記載のマルチチャネル情報処理方法。

5

1 1.

前記複数の動画像情報に対応する音声データの音量を、各距離情報中の距離の値に反比例するように設定することを特徴とする、請求項9に記載のマルチチャネル情報処理方法。

10

1 2.

カーソルの表示位置から見た各動画像情報の表示位置の方向に関する方向情報を生成する段階と、

前記複数の動画像情報に対応する音声データが、前記音声出力装置の音像空間における前記距離情報および方向情報に応じた音像位置に定位するように、前記音声出力装置に出力を行う段階と、  
をさらに備える、請求項9に記載のマルチチャネル情報処理方法。

1 3.

20 前記複数の動画像情報に対応する音声データに含まれる言葉を音声認識する段階と、

音声認識した言葉を文字情報に変換して前記表示装置に表示する段階と、  
をさらに備える請求項9に記載のマルチチャネル情報処理方法。

25 1 4.

インターネットに接続する段階と、

音声認識した言葉をキーワードとしてインターネット上に存在する関連ホームページを検索する段階と、

検索された関連ホームページを前記表示装置上に表示する段階と、

を備える、請求項 13 に記載のマルチチャネル情報処理方法。

15.

- 5 前記表示装置に表示される複数の動画像情報から所定のアルゴリズムに基づいて特定の動画像情報を選択する段階と、  
選択した動画像情報に対応する音声データを前記音声出力装置に出力する段階と、  
をさらに備える、請求項 9 に記載のマルチチャネル情報処理方法。

10 16.

- 複数の動画像情報を同時に表示装置に表示する際のマルチチャネル情報処理方法のプログラムであって、  
前記複数の動画像情報の表示装置上での表示位置を決定する段階と、  
決定した表示位置に基づいて前記複数の動画像情報を出力する段階と、  
15 前記入力デバイスから入力されるカーソル指示情報を受け付ける段階と、  
前記カーソル指示情報に基づいてカーソルを表示するカーソル位置情報を算出する段階と、  
前記カーソル位置情報に基づいてカーソル画像情報を生成する段階と、  
前記複数の動画像情報と前記カーソル画像情報とを合成して表示画像を生成し  
20 これを前記表示装置上に出力する段階と、  
前記複数の動画像情報の表示位置と前記カーソル位置情報との間の距離を算出し距離情報を生成する段階と、  
前記距離情報に基づいて複数の動画像情報に対応する音声データの音量を決定し音声出力装置に出力する段階と、  
25 を含むマルチチャネル情報処理方法をコンピュータに実行させるためのプログラム。

17.

複数の動画像情報を同時に表示装置に表示する際のマルチチャネル情報処理方

法のプログラムであって、

前記複数の動画像情報の表示装置上での表示位置を決定する段階と、

決定した表示位置に基づいて前記複数の動画像情報を出力する段階と、

前記入力デバイスから入力されるカーソル指示情報を受け付ける段階と、

- 5 前記カーソル指示情報に基づいてカーソルを表示するカーソル位置情報を算出する段階と、

前記カーソル位置情報に基づいてカーソル画像情報を生成する段階と、

前記複数の動画像情報と前記カーソル画像情報とを合成して表示画像を生成しこれを前記表示装置上に出力する段階と、

- 10 前記複数の動画像情報の表示位置と前記カーソル位置情報との間の距離を算出し距離情報を生成する段階と、

前記複数の動画像情報について生成された距離情報のうち距離の値が最も小さい動画像情報に対応する音声データの音量を予め定められた音量 $V_a$ に設定し、他の動画像情報に対応する音声データの音量を前記音量 $V_a$ よりも小さな予め定められた音量 $V_b$ に設定し音声出力装置に出力する段階と、

- 15 含むマルチチャネル情報処理方法をコンピュータに実行させるためのプログラム。

18.

- 20 複数の動画像情報を同時に表示装置に表示する際のマルチチャネル情報処理方法のプログラムであって、

前記複数の動画像情報の表示装置上での表示位置を決定する段階と、

決定した表示位置に基づいて前記複数の動画像情報を出力する段階と、

前記入力デバイスから入力されるカーソル指示情報を受け付ける段階と、

- 25 前記カーソル指示情報に基づいてカーソルを表示するカーソル位置情報を算出する段階と、

前記カーソル位置情報に基づいてカーソル画像情報を生成する段階と、

前記複数の動画像情報と前記カーソル画像情報とを合成して表示画像を生成しこれを前記表示装置上に出力する段階と、

前記複数の動画像情報の表示位置と前記カーソル位置情報との間の距離を算出し距離情報を生成する段階と、

前記複数の動画像情報に対応する音声データの音量を、各距離情報中の距離の値に反比例するように設定し音声出力装置に出力する段階と、

- 5   を含むマルチチャネル情報処理方法をコンピュータに実行させるためのプログラム。

## 19.

- 10   複数の動画像情報を同時に表示装置に表示する際のマルチチャネル情報処理方法のプログラムであって、

前記複数の動画像情報の表示装置上での表示位置を決定する段階と、

決定した表示位置に基づいて前記複数の動画像情報を出力する段階と、

前記入力デバイスから入力されるカーソル指示情報を受け付ける段階と、

- 15   前記カーソル指示情報に基づいてカーソルを表示するカーソル位置情報を算出する段階と、

前記カーソル位置情報に基づいてカーソル画像情報を生成する段階と、

前記複数の動画像情報と前記カーソル画像情報とを合成して表示画像を生成しこれを前記表示装置上に出力する段階と、

- 20   前記複数の動画像情報の表示位置と前記カーソル位置情報との間の距離を算出し距離情報を生成する段階と、

カーソルの表示位置から見た各動画像情報の表示位置の方向に関する方向情報を生成する段階と、

- 25   前記複数の動画像情報に対応する音声データが、前記音声出力装置の音像空間における前記距離情報および方向情報に応じた音像位置に定位するように、前記音声出力装置に出力を行う段階と、

を含むマルチチャネル情報処理方法をコンピュータに実行させるためのプログラム。

## 20.

複数の動画像情報を同時に表示装置に表示する際のマルチチャネル情報処理方法のプログラムであって、

前記複数の動画像情報の表示装置上での表示位置を決定する段階と、

決定した表示位置に基づいて前記複数の動画像情報を出力する段階と、

5 前記入力デバイスから入力されるカーソル指示情報を受け付ける段階と、

前記カーソル指示情報に基づいてカーソルを表示するカーソル位置情報を算出する段階と、

前記カーソル位置情報に基づいてカーソル画像情報を生成する段階と、

前記複数の動画像情報に対応する音声データに含まれる言葉を音声認識する段

10 階と、

音声認識した言葉を文字情報に変換してこれを出力する段階と、

前記複数の動画像情報、前記カーソル画像情報および前記文字情報とを合成して表示画像を生成しこれを前記表示装置上に出力する段階と、

を備えるマルチチャネル情報処理方法をコンピュータに実行させるためのプログ

15 ラム。

## 21.

複数の動画像情報を同時に表示装置に表示する際のマルチチャネル情報処理方法のプログラムであって、

20 前記複数の動画像情報の表示装置上での表示位置を決定する段階と、

決定した表示位置に基づいて前記複数の動画像情報を出力する段階と、

前記入力デバイスから入力されるカーソル指示情報を受け付ける段階と、

前記カーソル指示情報に基づいてカーソルを表示するカーソル位置情報を算出する段階と、

25 前記カーソル位置情報に基づいてカーソル画像情報を生成する段階と、

前記複数の動画像情報の表示位置と前記カーソル位置情報との間の距離を算出し距離情報を生成する段階と、

前記距離情報に基づいて複数の動画像情報のうちから特定の動画像情報を選択し、この動画像情報に対応する音声データを音声出力装置に出力する段階と、



- 前記音声出力装置に出力する音声データに含まれる言葉を音声認識する段階と、  
音声認識した言葉を文字情報に変換してこれを出力する段階と、  
前記複数の動画像情報、前記カーソル画像情報および前記文字情報とを合成して表示画像を生成しこれを前記表示装置上に出力する段階と、  
5   を備えるマルチチャネル情報処理方法をコンピュータに実行させるためのプログラム。

## 22.

- 複数の動画像情報を同時に表示装置に表示する際のマルチチャネル情報処理方法のプログラムであって、  
前記複数の動画像情報の表示装置上での表示位置を決定する段階と、  
決定した表示位置に基づいて前記複数の動画像情報を出力する段階と、  
前記入力デバイスから入力されるカーソル指示情報を受け付ける段階と、  
前記カーソル指示情報に基づいてカーソルを表示するカーソル位置情報を算出  
15   する段階と、  
前記カーソル位置情報に基づいてカーソル画像情報を生成する段階と、  
前記複数の動画像情報の表示位置と前記カーソル位置情報との間の距離を算出し距離情報を生成する段階と、  
前記距離情報に基づいて複数の動画像情報のうちから特定の動画像情報を選択  
20   し、この動画像情報に対応する音声データを音声出力装置に出力する段階と、  
前記音声出力装置に出力する音声データに含まれる言葉を音声認識する段階と、  
インターネットに接続する段階と、  
音声認識した言葉をキーワードとしてインターネット上に存在する関連ホームページを検索する段階と、  
25   前記複数の動画像情報、前記カーソル画像情報および前記検索された関連ホームページを合成して表示画像を生成しこれを前記表示装置上に出力する段階と、  
を備えるマルチチャネル情報処理方法をコンピュータに実行させるためのプログラム。